DOCUMENT RESUME

ED 394 096 CG 026 923

AUTHOR Dent, Harold E.
TITLE Everything You Thought Was True about IQ Testing, but Isn't: A Reaction to "The Bell Curve."
PUB DATE 12 Aug 95
NOTE 14p.; Paper presented at the Annual Meeting of the American Psychological Association (103rd, New York, NY, August 11-15, 1995).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS History; Intelligence Quotient; *Intelligence Tests; Psychological Testing; Psychology; *Racial Bias; Racial Discrimination; *Science History; Social History; *Test Bias
IDENTIFIERS Bell Curve (Herrnstein and Murray)

ABSTRACT

Rather than focus on the numerous flaws in the book "The Bell Curve" (Herrnstein & Murray), this discussion focuses on the racism and bigoted beliefs of the pioneers in the mental measurement movement in the United States--beliefs which provided the background and opportunity for the publication of the book. A significant amount of these historical attitudes still permeate theory and practice in the field of psychological testing today. The paper contends that the professional psychological community has been remiss in fulfilling its moral obligation to insure that the public has accurate information on issues where psychological expertise is relevant. Racial relations in the United States are precariously brittle; thus, it is critical that the professional psychological community change its laissez faire stance, assert its moral leadership, and use this opportunity to set the scientific record straight. The American Psychological Association must articulate state-of-the-art information on these issues and exercise its influence on public policy instead of allowing others such as adherents to "The Bell Curve" continue to fuel hate and racial bigotry. Contains 44 references. (JBJ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 394 096

Everything You Thought Was True About IQ Testing, But Isn't:

A Reaction to *The Bell Curve*

Harold E. Dent, Ph.D.
Center For Minority Special Education
Hampton University
Hampton, VA

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. DENT

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at a symposium: The "Bell Curve" Controversy, at the American Psychological Association meeting, New York, N.Y.

August 12, 1995

BEST COPY AVAILABLE

CG026923

Everything You Thought Was True About IQ Testing, But Isn't:
A Reaction *to The Bell Curve.*

Harold E. Dent, Ph.D.
Center For Minority Special Education
Hampton University

I. Greetings!

Thank you, Dr. Fairchild, for that introduction. I too want to thank Dr. Marie Root for organizing this session on *The Bell Curve* and particular thanks for inviting me to be a panel member among these distinguished psychologists. This session allows me the opportunity to say to the APA membership a few things that I have wanted to share in a forum such as this for a long, long time. I guess the point is, if you stick around long enough you just might get your opportunity. Again, thank you, Dr. Root.

II. History of racism in testing movement

My distinguished colleagues on this panel have identified numerous flaws in *The Bell Curve*, scientific flaws as well as flaws in scholarship. Yet, in spite of these flaws, the Herrnstein-Murray book has garnered an enormous amount of attention from the press, the TV media, the psychological community, and the general public in the year since its publication. Rather than focus my remarks on the book, I want to focus your attention on the history of racism in our profession which has provided the background and opportunity, if not sanction, for publication of the distortions and outright propaganda espoused in the pages of *The Bell Curve* under the guise of science.

From its very beginnings, the mental measurement movement in this country has been characterized by efforts to advance the theory of white intellectual superiority over non-whites. As psychologists we are all aware that Lewis Terman and his colleagues at Stanford University translated the Binet-Simon Intelligence Scale into English, and we are aware that after making minor modifications they renamed it the Stanford-Binet Intelligence Scale. We are also aware of the differences between Alfred Binet's concept of intelligence and Lewis Terman's views on intelligence, particularly differences with regard to the genetic transmissibility and immutability of intelligence. Terman's views on racial differences are seldom emphasized in the literature and rarely, if ever, discussed in the graduate courses where the theory and administration of the Stanford-Binet scale are taught. Terman's views on intelligence and racial differences were consistent with, if not influenced by, the ideology embodied in the eugenics movement of the day. His beliefs are clearly expressed in the following quote from one of his writings:

> "...their dullness seems to be racial, or at least inherent in the family stock from which they come. The fact that one meets this type with such frequency in Indians, Mexicans, and Negroes, suggests quite forcibly that the whole question of racial differences in mental traits will have to be taken up anew by experimental methods. This writer predicts that when this is done, there will be discovered enormously significant racial differences in general intelligence, differences which cannot be wiped out by any scheme of mental culture." (Terman, 1916, p. 92).

Many American psychologists of that era shared similar views on racial differences and intelligence (Goddard, 1917; Yerkes, 1921; Brigham, 1923), and they were extremely influential in molding public policy. More recently others in the profession have sought to revive the public's interest in racial differences around the issue of intelligence (Shuey, 1966; Jensen, 1969, 1980). *The Bell Curve* is merely the latest of these divisive efforts designed specifically to generate social tension around racial issues as a means of influencing public policy. Vontress (1992) observed that these efforts occur with consistency during periods of economic instability.

Dr. Henry Goddard, another of the early pioneers in the mental measurement movement, also contributed to the legacy of scientific racism in this field through his desire to preserve the nation from the scourge of the feebleminded. Dr. Goddard, a Professor of Psychology at Princeton University, was also Director of Research at the Vineland School for the Feebleminded in New Jersey. He was a strong advocate for the sterilization of the feebleminded and a staunch believer in the fledgling field of psychometrics. Shortly after the new intelligence test was developed, Goddard and his students administered the Stanford-Binet Intelligence Scale to a small sample of recent arrivals at the immigration center on Ellis Island. He reported that 87 % of the Russians, 83 % of the Jews, 80 % of the Hungarians, and, 79 % of the Italians entering this country were "feebleminded" (Goddard, 1917).

Entry into World War I by the United States provided a golden opportunity for the mental measurement movement to advance its technology and generate more grist for their scientific racist propaganda mill. Draftees had to be screened, classified, and trained for thousands of different military specialties. Dr. Robert Yerkes, a Harvard psychologist, headed the army testing program which developed and administered the Army Alpha and the Army Beta tests to approximately two million inductees. Although these tests were essentially screening tests, they were quickly dubbed intelligence tests. After WW I, Col. Yerkes edited a voluminous report published by the National Academy of Science (1921) summarizing data obtained from the army intelligence testing. One of the amazing discoveries revealed by this massive testing of draftees, many of whom were foreign born and had immigrated to this country, was that the longer one lived in the United States the more intelligent one seemed to become. Specifically, it was reported that foreign born draftees who lived in this country more than twenty years obtained higher scores on the Army Alpha Intelligence Test than draftees who lived in the USA from fifteen to twenty years. Those who lived here for ten to 15 years scored higher than those who lived here five to ten years, and so on and so on. Apparently those dedicated scientists involved in the construction of the Army Alpha did not consider the culture specific bias inherent in the instrument as a possible factor contributing to the consistency in this pattern of scores.

Carl Brigham's book, *A Study of American Intelligence* (1923) also reported data from the testing of draftees being inducted into the armed forces during World War I. Brigham's findings supposedly supported Goddard and Yerkes' claims. Brigham reported that foreign born draftees from Northern European countries such as England, Holland, Germany, and Norway were intellectually superior to those from Southern and Eastern European countries (e.g., Italy, Poland, Russia, and the Slavic countries). He stressed the notion that immigrants from Southern and Eastern European countries were less desirable immigrants than Northern Europeans because of their limited intellectual abilities. This pseudo scientific interpretation of data had a profound impact on public opinion, and particularly on politicians. So much so that Congress enacted the Johnson-Lodge Immigration Restriction Act of 1924 which established immigration quotas based on the immigration rates of 1890. The year 1890 was selected as a benchmark on which to base

immigration quotas because prior to that year Northern Europeans migrated to America in large numbers and after 1890 there was a marked increase in the numbers of Southern Europeans who entered this country. Brigham (1930) later recanted his position on this issue and stated that his conclusions were unfounded. Nevertheless, the impact of the law had a horrendous effect on thousands of refugees who were prevented from entering this country in their desperate effort to flee from Nazi Germany during the years immediately preceding World War II.

Audrey Shuey's book *the Testing of Negro Intelligence* (Shuey, 1966) and Arthur Jensen's *Harvard Educational Review* article, How Much Can We Boost IQ and Achievement?" (Jensen, 1969) were undisguised, overt efforts to eliminate government support for enrichment programs for poor, minority group children, specifically the Head Start programs. Jensen's 1969 conclusion was identical to the 1994 message propagated by Herrnstein and Murray that infusion of federal dollars will not overcome the cognitive disadvantage imposed by the limited genetic endowment reflected in the low IQ scores which poor and minority group children obtain. These are a few examples of the efforts by members of the psychological community to advance theories of racial superiority based on distortions and misinterpretations of so-called scientific data. For more extensive discussions of this subject the reader is referred to: Block & Dworkin, 1978; Chase, 1977; Ehrlich & Feldman, 1977; Gould, 1981; Guthrie, 1976; Kamin, 1974; Lawler, 1978; Mensh & Mensh, 1991.

III. Psychological community's response to *The Bell Curve*

Throughout its history the professional psychological community (specifically, the American Psychological Association) has taken a traditionally self serving, "good ole boy" approach to issues of this nature, particularly issues which directly impact minorities. Protecting the interests of its membership has always taken precedence over the profession's responsibility for safeguarding the interest of the public. Publication of *The Bell Curve* provides one more opportunity for the professional psychological community (APA) to set the record straight about the distortions and misinterpretations perpetrated in this book under the mantle of science. It is this writer's belief that the professional psychological community has a responsibility to provide the public with state-of-the-art information on matters where psychological expertise is relevant, especially at times such as this when psychological information is purposefully distorted and used deceptively as in *The Bell Curve*. The facts which were distorted and/or misinterpreted in *The Bell Curve* could quickly be corrected by accurate information coming from a prestigious source such as the American Psychological Association. They include, but are not limited to: bias in psychological testing; racism in the mental measurement movement (referred to in the preceding paragraphs); the lack of agreement among psychologists on a definition of intelligence; the facts about the heritability of intelligence; the meaning of correlation; and even the facts about effective interventions. As a pivotal point for such a state-of- the-art-discussion, I will briefly comment on the most fundamental of these issues-bias in psychological testing.

IV. Bias in psychological testing

Cultural bias in psychological testing has been a serious concern for African Americans from the days when W. E. B. Du Bois called attention to the inherent dangers of a test that pretended to measure innate human intelligence when IQ tests were first introduced into this country. He waged an incessant campaign of caution through his articles in the NAACP's monthly publication, T*he Crisis* magazine, which he edited from 1910 to 1934 (Du Bois, 1940).

While there is a plethora of literature reporting research on the bias in standardized testing (Gould, 1982; Guthrie 1976; Hilliard, 1995; Lawler, 1978; Mensh & Mensh, 1991); testimony of leaders in the test industry in a federal court acknowledging the presence of bias (Munday, 1978; Thorndike 1978); as well as the ruling of a Federal District Court (Larry P. v. Riles, 1979), which was upheld by two appellate panels (Larry P. v. Riles, 1984, 1986), the official position of the professional psychological community has been to deny bias in testing (APA, 1995).

For years both the test industry and the APA have ignored requests from minority psychologists (The Association of Black Psychologists-ABPsi) to address the problem of cultural bias in standardized tests (Dent & Williams, 1972). The industry does not publicize the fact that bias favoring females was eliminated from the original IQ test (Loewen, 1993) and that until 1972 females averaged higher scores on the SAT (Loewen, 1993; Rosser, 1989), so much so that it had to be revised. If revisions can be made to remove and/or reverse gender bias in psychological testing, it is logical to assume that the more sophisticated statistical techniques currently available could enable the industry to eliminate cultural bias against minorities or the bias favoring the dominant cultural group in existing tests.

The basis of arguments proclaiming or disclaiming the presence of cultural bias in standardized testing hinges on the difference in definitions employed by each of the adversary groups in this dispute to describe bias. Minority psychologists (ABPsi) and others who advocate that bias exists describe cultural bias in terms of the inherent bias in the content of test items, the bias which enters through the standardization process, and the validation procedures in test construction (Dent, 1976; Jones, 1987). In contrast, the testing industry and the APA describe bias in terms of bias which is determined through statistical manipulation of test scores. The profession acknowledges that administration and interpretation of tests represent possible sources of bias, but refer to this merely as the misuse of tests.

Content bias in IQ tests can be clearly demonstrated by examination of specific items used in the particular test under scrutiny. (The use of the one example from an early edition of an IQ test is provided here purely for emphasis. It must be kept in mind that despite claims that revisions have removed bias, the correlation coefficients between different early and current versions of these instruments remain substantial and significant): Cultural bias in test items must also be understood in context of the basic assumptions which must be met if test results are to be considered valid. The first assumption is that all who take the test have had similar experiences or opportunity to have common experiences. Examination of the content of items on intelligence tests will make it apparent that the assumption of commonalty of experience cannot be met, nor can the assumption of opportunity for commonalty of experience be satisfied. Economic, geographic, as well as social factors greatly influence the accessibility, availability and opportunity for all members of this society to share common experiences. To penalize minorities and those from poor backgrounds for not having access or opportunity to experience society or the environment as others have does not suggest objectivity or fairness in the testing process nor does it insure accuracy in measurement of ability.

To ask a child who was born and lived all his/her life on an island such as Hawaii, where the directional frame of reference is the sea, "Makai" and the mountains, "Mauka", In what direction does the sun set?, is to place that child at an experiential disadvantage. The child's natural response is "Makai", but the only acceptable response listed in the manual is, "in the West". Similarly, to ask that same Hawaiian child, What would you do if you saw a train

approaching a broken track?, is to place that child at a disadvantage. There are no trains in Hawaii!

The Educational Testing Service (ETS), the largest producers of standardized tests in the world, is aware of factors which depress the scores of minority group members and women who take standardized tests. Research conducted by ETS staff indicates that the demand for speed (Schmitt & Bleistein, 1987; Schmitt & Dorans, 1987; and Dorans, Schmitt, & Bleistein, 1988), the use of homographs (Bleistein & Wright, 1987; Schmitt & Bleistein, 1987; Schmitt, 1988, and O'Neill, McPeek, & Wild, 1990), and certain sentence structure confuses African Americans, Asian Americans and Hispanic Americans and depresses their scores (Rogers & Kulick, 1986; Schmitt & Bleistein 1987; and O'Neill, McPeek, & Wild, 1990). O'Neill, McPeek and Wild (1990) and Schmitt and Dorans (1988) found that subject matter content of particular interest to gender or ethnic groups yield scores which favor those groups.

This research emphasizes how word usage and language is integrally related to content bias. Another factor which can constitute a source of bias is the use of analogies in test items (Dorans, 1982; Rogers & Kulick, 1986; and Schmitt & Bleistein, 1987). Lack of understanding of the key word in an analogy poses serious problems for minority test takers. Lack of facility with the English language will impose a handicap on a child from a non-English speaking home where the parents did not complete high school when compared with the English language facility of a child from a home where both parents were English speaking college graduates. Vocabulary test items are based on word frequency counts which were taken in majority communities in years past. No effort has been taken by the test industry to determine the frequency of word usage in minority communities. Yet, minority test takers are measured by the same vocabulary tests as majority test takers.

V. Standardization as a source of bias

The standardization or norming process is another source of bias which the test industry and the profession choose to ignore. In the construction of a standardized test, each item is administered to a try-out sample, which like a normative sample is representative of the total population. The responses of these small try-out samples determine which items will be selected for inclusion in the test when completed. Minority groups are represented in these try-out samples and in the normative samples in the same proportions as they are in the population. They (minority individuals) do not cluster in large enough numbers at any point in the distribution in either the try-out samples or the normative samples to have any influence on the outcome of the selection of the items or the norms established for the test. David Wechsler expressed caution about mixing ethnic groups in the standardization sample in his first book, *The Measurement of Intelligence* (1944).

> "[We] have eliminated the colored vs. White factor by admitting at the outset that our norms cannot be used for the colored population of the United States. Though we have tested a large number of colored persons, our standardization is based upon White persons only. We omitted the colored population from our first standardization because we did not feel that norms derived by mixing the population could be interpreted without special provisos and reservations" (p. 107).

Green made a similar observation in a monograph entitled, *Racial and Ethnic Bias in Test Construction* (1972). He states:

> "Just as the degree of minority representation in standardization samples can have only a small influence on the norms, minority presence in tryout samples dominated by some solid majority will not accomplish much." (p. 14).

In recent years, particularly since the passage of P.L. 94-142 (The Education for All Handicapped Childrens Act) increased attention has been focused on the issue of test bias and non-discriminatory testing. Test producers have tried to convey the impression that including minorities in the standardization samples render tests free of bias, and have made a point to publicize that minorities were included in the test norms. Despite these marketing ploys, there have been no data reported in the literature to cast doubt on the conclusions drawn by Wechsler and Green. In other words, the message we all received in Psych. 101 still holds true, that a test should be applied only to the population for which it was designed.

In direct response to the demand by the newly formed Association of Black Psychologists to declare a moratorium on the use of culturally biased IQ tests on African American children, and as part of a continuing effort to justify the practice of applying standardized tests to minority groups, the American Psychological Association appointed a blue ribbon committee in the early 1970's to delineate the conditions for the use of psychological and educational tests with minority group children in schools. This blue ribbon committee was composed of experts in psychometrics, whose resulting report, *The Educational Use of Tests With Disadvantaged Students*, (Cleary, Humphreys, Kendrick, & Wesman, 1975), was viewed by many as the definitive solution to the controversy associated with the application of standardized IQ tests to African American children. The report outlined a set of conditions under which a test could be considered "fair" for a particular use with separate groups of examinees. This report was cited frequently by the defense in the litigation challenging the use of IQ tests on African American students in California, (Larry P. vs. Riles, 1979). In essence, the Cleary, et. al., report stated that a test was fair and could be used with different populations if three conditions were met. Those conditions were:

1. The regression lines of the distributions of scores (on the same standardized test) of different groups were parallel;
2. The slope of the regression lines of these separate distributions was similar; and,
3. The correlation between the criteria and the test scores were similar for the two groups (see Figure 1).

Figure 1. goes here

Mercer (1979) applied these criteria to a set of data obtained in a statewide study and testified that: 1) the correlation coefficients between IQ test scores (Verbal Scale of the WISC) and grade point averages (GPAs) for African American students and white students, in grades kindergarten through sixth, were significantly different; 2) the regression lines for the two distributions were not parallel; and, 3) the slope of the regression lines of the two distributions were not similar. In fact the regression lines intersected. Figure 2 represents the superimposed regression lines for the two distributions of IQ scores and GPAs of the two groups. The correlation coefficient between these two variables (IQ scores and GPAs) for African American students is $r = .20$ and for white students it is $r = .458$. The index of the slope of the regression

line for the scores of African American students is .77, whereas the index of the slope of the regression line for the distribution of white students scores is 2.31.

Figure 2. goes here

The only logical conclusion one can draw from these data is that the WISC does not meet the APA criteria of fairness for application to African American elementary school age children. A number of articles critical of the Larry P. decision have appeared in the literature (Lambert, 1981; Sattler, 1981; Prasse & Reschly, 1986; Taylor, 1990; and Elliott, 1992), but none of the authors valued this extremely impressive and relevant data as important enough to cite.

VI. Federal Statutes

The American Psychological Association, local school districts across the nation, and the Federal Government have virtually ignored the federal laws which Judge Robert F. Peckham found had been violated when he issued his landmark decision in the Larry P. case. Judge Peckham found that IQ tests had not been validated for the specific purpose for which they were used, which violated Section 504 of The Rehabilitation Act of 1973; and, that IQ tests discriminated against African American children because they did not account for the background and experience of these children, which violated P.L.94-142, The Education of All Handicapped Children's Act. Judge Peckham's decision has been reviewed and upheld by two different Federal Appellate Court panels (Larry P. v. Riles, 1979, 1984, 1986). However, these tests are still being used on African American students by psychologists in school districts throughout the country on a daily basis with impunity and with the tacit sanction of the APA. It should be emphasized here that the APA has been conspicuously silent on the issue of the violation of federal law in the application of IQ tests on minority group children. Yet in the most recent public defense of IQ tests, "Intelligence: Knowns and Unknowns"(APA, 1995) which was called a response to *The Bell Curve*, the traditional position of the APA was reiterated once again; the cause of the 15 point differential between IQ scores of African Americans and whites, "...is not known; it is apparently not due to any simple form of bias in the content or the administration of the test themselves."(P. 43). Of interest is the fact that the most important legal decision in the past two decades involving intelligence tests and race, Larry P. v. Riles (1979), was not mentioned in this latest APA document on intelligence. To this writer, such action is indicative of APA's blatant willingness to ignore critical factors which support minority psychologists' and the minority community's position on bias in standardized testing.

Critics expressing their negative opinions about the Larry P. decision often compare the Larry P. v. Riles (1979) and the P.A.S.E. v. Hannon (1980) cases, simply because they dealt with similar issues, the cultural bias in IQ tests. These writers want their readers to believe that two federal judges presented with identical information rendered completely opposite decisions. Analysis of some very basic facts about these cases will indicate how misleading such reports are. The Larry P. trial extended over a period of eight months, involved the testimony of more than two dozen expert witnesses in the fields of psychometrics, psychological testing and test construction, and required over ten thousand pages of transcript to record. Whereas, the P.A.S.E. trial was completed in two weeks. These critics fail to mention that Judge Peckham carefully weighed the testimony of the expert witnesses for both sides and cited reasons for accepting and/or rejecting their respective positions in his decision. Whereas Judge Grady stated

that because the expert witnesses for both sides in the P.A.S.E. case could not agree he had set himself up as the sole determinant of cultural bias in test item content. He actually read every item of the WISC-R and the Stanford-Binet Intelligence Scale into the court record and decided (by some undisclosed intuitive process) which items were biased and which items were not biased. He concluded that only a small number of items were biased and ruled that was not enough to make the tests biased against the African American student plaintiffs. Critics of Larry P. also fail to mention that in an article comparing these two decisions, the attorney for the APA, Donald Boers, himself a psychologist, offered a very unflattering comment about Judge Grady's method of determining cultural bias. Boers stated, "The method by which Judge Grady reached that judgment is embarrassingly unsophisticated and ingenuous."(p.1049, 1981). Those who profess that these decisions balance each other should ask themselves if they would prefer to have two weeks or eight months to present their case, and if they would prefer to have the tribunal dismiss their experts' testimony because the opinions of opposing experts differed.

In summary, rather than focus on the myriad of flaws, misinterpretations, and distortions replete in *The Bell Curve*, this discussion shifted the readers attention to the history of racism and bigoted beliefs of the pioneers in the mental measurement movement in this country much of which still permeates theory and practice in the field of psychological testing today. This writer contends that the professional psychological community has been woefully remiss in fulfilling its moral obligation to insure that the public has accurate information on issues where psychological expertise is relevant. By its silence on issues such as race and gender bias in testing, the professional psychological community has given tacit sanction to the authors of *The Bell Curve* and others whose true agenda is the dissemination of "politically correct" ideologies of the day. This writer believes it is critical at this time when racial tensions in this country are precariously brittle that the professional psychological community change its laissez faire stance and assert its moral leadership and use this opportunity to set the scientific record straight. The APA must articulate state-of-the-art information on these issues and exercise its influence on public policy instead of allowing others such as the authors of *The Bell Curve* continue to fuel the flames of hate and racial bigotry with their propaganda.

BIBLIOGRAPHY

American Psychological Association, (1995). *Intelligence: Knowns and Unknowns*. Report of a Task Force established by the Board of Scientific Affairs of the American Psychological Association.

Bersoff, D. (1981). Testing and the Law. *American Psychologist*. Vol. 36, No.10, 1047-1056.

Bleistein, C. A. & Wright, D. (1987). Assessment of unexpected differential item difficulty for Asian American candidates in the Scholastic Aptitude Test. In A. P. Schmitt & N. J. Dorans (Eds.) *Differential item functioning on the Scholastic Aptitude Test*. (RM-87-1). Princeton, NJ: Educational Testing Service.

Block, N. J., & Dworkin, G. (1976). *The IQ controversy: Critical readings*. New York: Pantheon.

Brigham, C. C. (1923). *A study of American intelligence*. Princeton, NJ: Princeton University Press.

Brigham, C. C. (1930). Intelligence testing of immigrant groups. *Psychological Review*. 37:158-165.

Chase, A. (1977). *The Legacy of Malthus: The social costs of the new scientific racism*. New York: Knopf.

Cleary, T. A., Humphreys, L.G. Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist*. January:15-41.

Dent, H. E. (1976). Assessing Black children for mainstream placement. In R.L. Jones (Ed.), *Mainstreaming and the minority child*. Reston, VA: The Council for Exceptional Children.

Dent, H. E. & Williams, R. L. (1972, August). *Psychological testing: A position paper of the Association of Black Psychologists*. Paper presented at the American Psychological Association annual convention, Montreal, Canada.

Du Bois, W. E. B. (1940). *Dusk of dawn: An essay toward an autobiography of a race concept*. With a tribute to Dr. Du Bois by Martin Luther King, Jr. New York: Shocken Books, 1968. First published in 1940.

Ehrlich, P. R., & Feldman, S. S. (1977). *The race comb: Skin color, prejudice, and intelligence*. New York: Quadrangle.

Elliott, R. (1992). Larry P., P.A.S.E., and social science in the courtroom: The science and politics of identifying and educating very slow learners. In Haywood, H. C. & Tzuriel, D. (Eds.) *Interactive assessment*. New York: Springer-Verlag.

Goddard, H. H. (1917). Mental tests and the immigrant. *Journal of Delinquency* 2:243-277.

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton Press.

Green, D. R. (1972). *Racial and ethnic bias in test construction*. (Contract No. OEC9-70-0058 [57]). Final report of the United States Office of Education. Washington, DC.

Guthrie, R. V. (1976). *Even the rat was white: A historical view of psychology*. New York: Harper & Row.

Hilliard, A. G. (1995, August). *Either a paradigm shift or no mental measurement*. Paper presented to the American Psychological Association at the Annual Meeting, New York.

Jensen, A. (1969). How much can we boost IQ and achievement? *Harvard Educational Review*, 39,81.

Jensen, A. (1980). *Bias in mental testing*. New York: Free Press.

Jones, R. L. (Ed.). (1976). *Mainstreaming and the minority child*. Reston, VA: The Council for Exceptional children.

Kamin, L. J. (1978). *The science and politics of IQ*.. Potomac, MD: Erlbaum.

Lambert, N. (1981). Psychological evidence in Larry P. v. Wilson Riles: An evaluation for the defense. *American Psychologist*, 36, 937-952.

Larry P. et al., v. Wilson Riles et al. C-71-2270 RFP, (ND Cal. 1979), Aff'd 793 F. 2nd 979 (9th Cir. January 23, 1984, as Amended, June 25,1986).

Lawler, J. M. (1978). *IQ, heritability and racism*. New York: International.

Loewen, J. W. (1993). A sociological view of aptitude tests. *The validity of testing in education and employment*.(Report No. TM 021 009) Washington, D.C., U. S. Commission on Civil Rights. (ERIC Document Reproduction Service No. ED 366 625)

Mensh, E., & Mensh, H. (1991). *The IQ mythology: Class, race, gender, and inequality*. Carbondale, IL: Southern Illinois University Press.

Mercer, J. R. (1979). Expert testimony in the Larry P. trial (Larry P. v. Riles. 495 F. Supp., ND. Cal.).

Munday, L. (1979). Expert testimony in the Larry P. trial (Larry P. v. Riles. 495 F. Supp., ND. Cal.).

O'Neill, K. A., McPeek, M. W., & Wild, C. L. (1990). *Characteristics of quantitative items in graduate admissions tests that show differential items functioning for men and women.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, Boston, Massachusetts.

P.A.S.E. (People in Action on Special Education) v. Joseph P. Hannon. (1980). 506 F. Supp., 831. U.S. District Court, ND. Ill, Eastern Division, No.74.

Prasse, D. & Reschly, D. (1986). Larry P.: Case of segregation, testing, or program efficacy? *Exceptional Children*, 52, 333-346.

Rogers, J. & Kulich, E. (1986, April). *An investigation of unexpected differences in item performance between Blacks and Whites taking the SAT.* Paper presented at the annual meeting of the National Council on Measurement in Education. San Francisco, CA.

Rosser, P. (1989). *The SAT Gender Gap: Identifying the Causes*. Washington, D.C.: Center for Women Policy Studies.

Sattler, J. M.(1981). Intelligence tests on trial: An "interview" with judges Robert F. Peckham and John F. Grady. *Journal of Social Psychology*, 19, 359-369.

Schmitt, A. P., & Bleistein, C. A. (1987). *Factors affecting differential item functioning for Black examinees on Scholastic Aptitude Test analogy items.* (RR-87-23). Princeton, NJ: Educational Testing Service.

Schmitt, A. P. & Dorans, N. J. (1988). *Differential item functioning for minority examinees on the SAT.* (RR-88-32) Princeton, NJ: Educational Testing Service.

Shuey, A. (1966). *The testing of Negro Intelligence*. 2nd. ed. New York, NY: Social Science Press.

Taylor, R. L. (1990). The Larry P. decision a decade later: Problems and future directions. *Mental Retardation*, iii-v.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale.* Boston, MA: Houghton-Mifflin.

Thorndike, R. (1979). Expert testimony in the Larry P. trial. (Larry P. v. Riles. 495 F. Supp., ND. Cal.).

Vontress, C. E. (1992). Hard times and meanness. *Black Issues in Higher Education.* 8, (22), January 2.

Wechsler, D. (1944). *The measurement of intelligence*. Baltimore, MD: Williams and Williams.

Yerkes, R. M., (Ed). (1921). Psychological examining in the United States Army. *Memoirs of the National Academy of Sciences*. vol. 15. Washington, DC: U.S. Government Printing Office.